

Pseudo-Time Methods for Constrained Optimization Problems Governed by PDE

Shlomo Ta'asan*
Carnegie Mellon University
and
Institute for Computer Applications in Science and Engineering

Abstract

In this paper we present a novel method for solving optimization problems governed by partial differential equations. Existing methods use gradient information in marching toward the minimum, where the constrained PDE is solved once (sometimes only approximately) per each optimization step. Such methods can be viewed as a marching techniques on the intersection of the state and costate hypersurfaces while improving the residuals of the design equation per each iteration. In contrast, the method presented here march on the design hypersurface and at each iteration improve the residuals of the state and costate equations. The new method is usually much less expensive per iteration step, since in most problems of practical interest the design equation involves much fewer unknowns than either the state or costate equations. Convergence is shown using energy estimates for the evolution equations governing the iterative process. Numerical tests shows that the new method allows the solution of the optimization problem in cost equivalent to solving the analysis problem just a few times, independent of the number of design parameters. The method can be applied using single grid iterations as well as with multigrid solvers.

*This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-19480 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681-0001.

1 Introduction

In the last few years there has been a growing interest in the numerical solution of optimization problems governed by large scale problems. This new interest is a direct result of the improvement in computer technology. Probably the most challenging problems are those which involve complex governing equations such as the Euler, Navier-Stokes, Acoustic wave, and Maxwell's. Some global quantities governed by the solutions of such equations are required to be minimized (maximized) in terms of some prescribed design variables. The resulting mathematical problem is formulated as a constrained optimization problem which can sometimes be viewed as a control problem. Most existing algorithms use gradient information for reaching the minimum, possibly together with preconditioners for accelerating convergence.

Efficient gradient calculation can be done using the adjoint equations, and in area of aerodynamics design, this approach was first suggested in [J]. There the steepest descent method was employed there and the adjoint equations were used for efficient calculation of gradients. In this approach, each optimization step requires the solution of the state and costate equations and an efficient implementation is achieved by using multigrid methods for both equations. No acceleration of the optimization process was involved in this work.

The one shot method proposed in [T1] for control problems, also uses the adjoint method together with multigrid acceleration for state and costate, but also include an acceleration of the minimization process. Its development so far has been for problems with elliptic partial differential equations as constraints. The main idea is that smooth perturbations in the data of the problem introduce smooth changes in the solution, and highly oscillatory changes in the data produces highly oscillatory changes in the solution. Moreover, highly oscillatory changes are localized. These observations enable the construction of very efficient optimization procedure, in addition to very efficient solvers for the state and costate variables. Design variables that correspond to smooth changes in the solution are solved for on coarse levels and those corresponding to highly oscillatory changes are solved for on appropriate finer grids. The resulting method can be viewed as a preconditioning of the gradient descent method where the new condition number is independent of the grid size, and is of order 1. Thus, within a few optimization iterations one reaches the minimum. The method was first developed for a small dimensional parameter space, where the optimization was done on the coarsest grid, yet converging to the fine grid solution [T1]. Later in [TKS1], [TKS2] the method was applied to cases of a moderate number of design variables and where the optimization was done on few of the coarsest grids. The extension of these ideas to the infinite dimensional parameter space was done in [AT1],[AT2] where both boundary control as well as shape design problems were treated. In [AT1],[AT2] an important new analysis for the structure of the functional near the minimum was introduced. That analysis also enables the construction of efficient relaxation for multigrid methods and preconditioners for single grid techniques. Moreover, it can give essential information about the structure of the minimum including the condition number for the optimization problem, the well-posedness (ill-posedness) of the problem, and can

suggest appropriate regularization techniques. Experiments with the one shot method for finite dimensional and infinite dimensional design spaces showed that the convergence rate is practically independent of the number of design parameters.

The necessity of using multigrid algorithms in the one shot methods is certainly a disadvantage since in many engineering applications the underlying solvers do not use multigrid methods. This drawback has led to inquiries in other directions, but still aiming at algorithms that solve the full optimization problem in one shot, i.e., in a cost not much larger than that of solving the analysis problem.

The first observation made was that the solution when using the adjoint method is an intersection point of three hypersurfaces describing the state equation, costate-state equation and design equation (together forming the necessary conditions of the minimization problem). The adjoint method can be viewed as marching on the intersection of the hypersurfaces corresponding to state and costate variables, in the direction of the intersection with the design hypersurface. Since in most applications the number of design variables is significantly smaller than the number of state or costate variables, marching in the design hypersurface is a much less expensive process than the adjoint method, and may serve as a solution process for the optimization problem.

Methods that march on the design hypersurface are not based on gradients and their convergence properties are different. In this paper we construct and analyze methods of this type by embedding the necessary conditions into an evolution equation so that the solution evolves in the design hypersurface. Energy estimates are used to prove convergence.

The new methods which are stable approximations to evolution processes can be accelerated using multigrid or other acceleration techniques. Numerical results for model problems are presented and demonstrate the effectiveness of the method. It is shown that the full optimization problem is solved in a computer time equivalent to just a few solutions of the analysis problem. The method seems to converge in a rate independent of the number of design variables.

2 On Adjoint Methods

We consider the following constrained minimization problem

$$\min_u E(u, \phi(u)) \tag{1}$$

$$L(u, \phi) = 0 \tag{2}$$

where $L(u, \cdot)$ is a partial differential operator (possibly nonlinear) defined on a Hilbert space \mathcal{H} of functions defined on a domain Ω . The design variable is assumed to be in some other Hilbert space \mathcal{U} , for example, functions defined on the boundary $\partial\Omega$, or part of it.

The (formal) necessary conditions are

$$\begin{aligned} L(u, \phi) &= 0 && \text{State Equation} \\ L_\phi^* \lambda + E_\phi &= 0 && \text{Costate Equation} \\ L_u^* \lambda + E_u &= 0 && \text{Design Equation} \end{aligned} \tag{3}$$

and we assume the existence of solutions for both the state and costate equations.

It can be shown that the gradient of $E(u, \phi(u))$ is given by

$$\mathcal{A}(u) = L_u^* \lambda(\phi, u) + E_u(u, \phi(u)) \quad (4)$$

where $\phi(u), \lambda(\phi, u)$ are the solution of the state and costate equations. The quantity $-\mathcal{A}(u)$ can serve as a minimization direction (steepest descent).

The adjoint method consists of solving the state and costate equations at each update step of the design variables. Thus it can be viewed as an approximation to the following evolution process.

$$L(u, \phi) = 0 \quad (5)$$

$$L_\phi^* \lambda + E_\phi = 0 \quad (6)$$

$$\frac{d}{dt}u + L_u^* \lambda + E_u = 0 \quad (7)$$

where $\frac{d}{dt}u$ represent the derivative of u with respect to the pseudo-time variable introduced into the problem. The actual iteration method is obtained by replacing $\frac{d}{dt}u$ with $(u^{n+1} - u^n)/\delta t$, for a sufficiently small δt

The full algorithm can be viewed as a solver for the equation

$$\mathcal{A}(u) = 0 \quad (8)$$

for the variable u . A crucial quantity to consider for analyzing the efficiency of different algorithms is the mapping

$$u \longrightarrow \mathcal{A}(u) \quad (9)$$

For problems arising from partial differential equation this mapping is a differential or a pseudo-differential operator and bad conditioning is anticipated. Preconditioning of basic iterative methods such as the steepest descent, is needed.

The one shot methods [T1],[TKS1],[AT1],[AT2] were aiming at a preconditioning of the gradient algorithm in such a way that an order one condition number is obtained. In such a case the number of minimization step required to reach the minimum is independent of the size of the problem, i.e., the number of unknowns for u . This approach was found to be very successful for elliptic equations. The idea is to exploit the locality of high frequencies in the algorithm, as well as the fact that high frequencies in the design variables are related to high frequencies of the state variable and vice versa. Finite and infinite dimensional design spaces have been considered with application to aerodynamics problems, and other shape design problems.

Another possible direction, which was not explored, is to construct single grid preconditioners based on the form of the symbol of the operator \mathcal{A} . This idea will be discussed elsewhere.

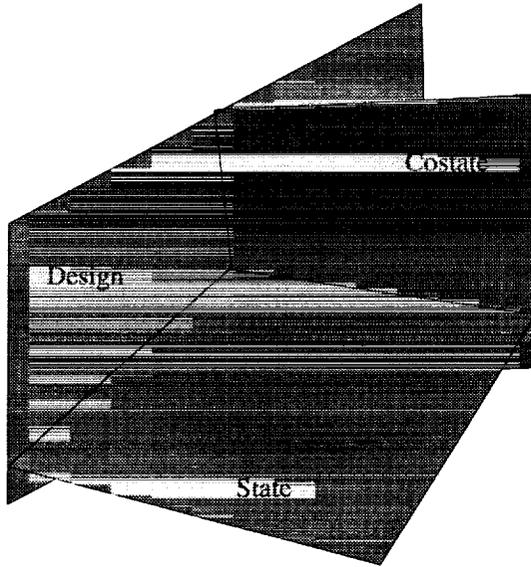


Figure 1: Hypersurfaces for state, costate and design equations

3 The New Approach

The solution of the minimization problem is the intersection point of the hypersurfaces defined by state, costate, and design equations, see figure 1. Gradient descent methods for constrained optimization problems march along the intersection of the state and costate hypersurfaces. Each step in such a process requires the solution of two large scale problems, namely, the discretized PDEs. Since in many applications the number of unknowns in either the state or costate equations is significantly larger than that in the design equation, marching on the hypersurface defined by the design equation is a much less expensive process than that of marching along the state and the costate hypersurfaces. This is the main idea of the new approach.

Each step in the minimization algorithms presented here improve the solution of the state and costate equations, for example, by improving the distance to the hypersurfaces defined by the state and costate equations. In addition each step is such that the approximate solution lies on the design hypersurface.

The construction of algorithms that march along the design hypersurface and converge to the minimum of the optimization problem can be done for a wide class of problems governed by PDE. The approach taken here is to look at iterative methods for the solution of the state and costate equations as a stable approximation to the evolution equations governed by the constrained PDE. The construction of the method is done in two steps. In the first the stationary PDE (the necessary conditions) is embedded into an evolution PDE for which the solution evolves in the design hyper-

surface, and an energy estimate ensuring convergence is derived. The second step involves a stable and consistent discretization of the pseudo-time dependent problem which is usually straightforward.

A technical difficulty which needs some explanation is related to the problem of staying on the design hypersurface. Assume that we are given an iterative method for calculating the solutions of the state and costate equations. Let the change produced in ϕ and λ be $\tilde{\phi}$ and $\tilde{\lambda}$ respectively. In order to remain on the design hypersurface it is necessary to calculate a change in u , namely, \tilde{u} such that

$$\mathcal{A}(u + \tilde{u}, \phi + \tilde{\phi}, \lambda + \tilde{\lambda}) = 0 \quad (10)$$

An approximation to this equation is

$$\frac{\partial \mathcal{A}}{\partial u} \tilde{u} = -\frac{\partial \mathcal{A}}{\partial \phi} \tilde{\phi} - \frac{\partial \mathcal{A}}{\partial \lambda} \tilde{\lambda} \quad (11)$$

This representation is useful when $\frac{\partial \mathcal{A}}{\partial u}$ is an invertible operator. Note that the solution of this equation involves a system whose size is identical to that of the number of design variables, which is significantly smaller than that of the state or the costate equations. While the operator $\frac{\partial \mathcal{A}}{\partial u} + \frac{\partial \mathcal{A}}{\partial \phi} \phi_u + \frac{\partial \mathcal{A}}{\partial \lambda} \lambda_u$ is invertible, it is not convenient to work with; however $\mathcal{A}_u = \frac{\partial \mathcal{A}}{\partial u}$ is simple and easy to manipulate.

In practice, \mathcal{A}_u may not be invertible and the update of the u requires a different process. In problems arising from partial differential equations in which the design variables are defined on the boundary only, the design equation is an additional boundary condition for the system, for the extra unknowns, namely, the design variables. In that case per each iteration step of the method presented here require the simultaneous solution of the three boundary conditions for the state equation, the costate equation and the design equation. These three conditions together involve only a fraction more work than that of the boundary conditions for the state and costate equations. In cases where the set of the three boundary conditions cannot be solved for the boundary state, costate, and design variables, one should include equations from the interior. This is a typical case when considering systems of partial differential equations in several dimensions.

In case that the linearized operator L_ϕ is coercive and the design equation can be solved for the design variables, keeping the state and costate variables fixed, one can view the method presented here as an approximation to the following time dependent problem

$$\frac{d}{dt} \phi + L(u, \phi) = 0 \quad (12)$$

$$\frac{d}{dt} \lambda + L_\phi^T \lambda + E_\phi = 0 \quad (13)$$

$$L_u \lambda + E_u = 0 \quad (14)$$

where the last equation is essentially an extra boundary condition for the design variables.

4 Examples

In this section we show a few examples of using the idea outlined in the previous section. We prove an energy estimate for each of the examples, ensuring convergence.

Example I: Distributed Control Let $\Omega \subset \mathbb{R}^n$ and consider the optimization problem

$$\min_u \frac{1}{2} \int_{\Omega} (\phi - \phi^*)^2 dx + \frac{1}{2} \sigma \int_{\Omega} u^2 \quad (15)$$

subject to

$$\begin{cases} \Delta \phi = u & \Omega \\ \phi = 0 & \partial\Omega \end{cases} \quad (16)$$

The necessary conditions are

$$\begin{cases} \Delta \phi = u & \Omega \\ \Delta \lambda + \phi = \phi^* & \Omega \\ \sigma u - \lambda = 0 & \Omega \\ \phi = 0 & \partial\Omega \\ \lambda = 0 & \partial\Omega \end{cases} \quad (17)$$

Consider the pseudo-time embedding

$$\begin{aligned} \frac{d}{dt} \phi &= \Delta \phi - u & \Omega \\ \frac{d}{dt} \lambda &= \Delta \lambda + \phi - \phi^* & \Omega \\ \sigma u - \lambda &= 0 & \Omega \\ \phi &= 0 & \partial\Omega \\ \lambda &= 0 & \partial\Omega \end{aligned} \quad (18)$$

We show that the error term in ϕ, λ, u , tend to zero as t approaches infinity. These error terms satisfy the same equations as their corresponding quantities ϕ, λ, u but with zero source terms. So without loss of generality we consider our problem with $\phi^* = 0$.

The proof uses Poincare's inequality in the form

$$\|\psi\|^2 \leq C(\|\nabla \psi\|^2 + (\int_{\Gamma} \phi ds)^2) \quad C > 0 \quad (19)$$

where $\Gamma \subset \partial\Omega$, and $C > 0$ is a constant independent of $\psi \in H^1(\Omega)$. The norm used above and in the rest of the paper is the L_2 norm on Ω . The use of this theorem will be for functions vanishing on part of the boundary denoted by Γ .

For this example we take $\Gamma = \partial\Omega$ since the boundary condition for the errors in both ϕ and λ is zero on $\partial\Omega$. Multiplying the first two equations in (17) by ϕ and λ respectively we get using integration by parts and the Poincare inequality

$$\frac{1}{2} \frac{d}{dt} (\sigma \|\phi\|^2 + \|\lambda\|^2) = -\sigma \|\nabla \phi\|^2 - \|\nabla \lambda\|^2 \leq -C(\sigma \|\phi\|^2 + \|\lambda\|^2) \quad (20)$$

for some constant $C > 0$, independent of ϕ, λ .

This implies that $\sigma\|\phi\|^2 + \|\lambda\|^2$ decay exponentially with rate $\exp(-Ct)$. That is, the pseudo-time embedding converges to the minimum, at a rate determined by the constant C .

Example II: Boundary Control

The next example is of a boundary control. Let $\Omega \subset \mathbb{R}^n$, $\Gamma_1 \cup \Gamma_2 = \partial\Omega$, $\Gamma_1 \cap \Gamma_2 = \emptyset$ and consider

$$\min_u \frac{1}{2} \int_{\partial\Gamma_1} (\phi - \phi^*)^2 dx + \frac{1}{2} \sigma \int_{\partial\Gamma_1} u^2 \quad (21)$$

subject to

$$\begin{cases} \Delta\phi = 0 & \Omega \\ \frac{\partial\phi}{\partial n} = u & \Gamma_1 \\ \phi = 0 & \Gamma_2 \end{cases} \quad (22)$$

The necessary conditions are

$$\begin{cases} \Delta\phi = 0 & \Omega \\ \Delta\lambda = 0 & \Omega \\ \frac{\partial\phi}{\partial n} = u & \partial\Omega \\ -\frac{\partial\lambda}{\partial n} + \phi = \phi^* & \Gamma_1 \\ \sigma u + \lambda = 0 & \Gamma_1 \\ \phi = 0 & \Gamma_2 \\ \lambda = 0 & \Gamma_2 \end{cases} \quad (23)$$

The time embedding used for this problem is

$$\begin{cases} \frac{d}{dt}\phi = \Delta\phi & \Omega \\ \frac{d}{dt}\lambda = \Delta\lambda & \Omega \\ \frac{\partial\phi}{\partial n} = u & \partial\Omega \\ -\frac{\partial\lambda}{\partial n} + \phi = \phi^* & \Gamma_1 \\ \sigma u + \lambda = 0 & \Gamma_1 \\ \phi = 0 & \Gamma_2 \\ \lambda = 0 & \Gamma_2 \end{cases} \quad (24)$$

In this case the use of Poincare's inequality is done for $\Gamma = \Gamma_2$. Similarly to the previous example it can be shown that

$$\frac{1}{2} \frac{d}{dt} (\sigma\|\phi\|^2 + \|\lambda\|^2) = -\sigma\|\nabla\phi\|^2 - \|\nabla\lambda\|^2 \leq -C(\sigma\|\phi\|^2 + \|\lambda\|^2) \quad (25)$$

with a different constant than that of example I. Again this estimate implies the exponential decay of errors. Thus, convergence of ϕ and λ is ensured, and therefore also of u from the Neumann boundary condition for ϕ .

Example III: System of First Order

Let $x \in \mathbb{R}^d$, A_1, \dots, A_d be symmetric constant matrices, $\phi = (\phi_1, \dots, \phi_n)$ defined on $\Omega \subset \mathbb{R}^d$. We introduce the decomposition of $\phi|_{\partial\Omega} = (\phi_+, \phi_0, \phi_-)$ as follows. Let $\mathbf{A} = (A_1, \dots, A_d)$ and \mathbf{n} be the outward normal to the boundary $\partial\Omega$. The matrix $\mathbf{A} \cdot \mathbf{n}$ is symmetric and therefore has real eigenvalues and a complete set of eigenvectors. Let $\phi = (\phi_-, \phi_0, \phi_+)$ be a decomposition into the direct sum of the subspaces of $\mathbf{A} \cdot \mathbf{n}$ corresponding to negative, zero and positive eigenvalues. For simplicity we also assume that $\mathbf{A} \cdot \mathbf{n}$ has zero eigenvalues on isolated points on the boundary $\partial\Omega$. Consider the following problem

$$\min_u \frac{1}{2} \int_{\Gamma_1} (\phi_- - g)^2 ds + \frac{1}{2} \sigma \int_{\Gamma_1} u^2 ds \quad (26)$$

where ϕ is the solution of

$$\begin{aligned} \sum_{j=1}^d A_j \frac{\partial \phi}{\partial x_j} &= 0 & \Omega \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= u & \Gamma_1 \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= 0 & \Gamma_2 \end{aligned} \quad (27)$$

where $\Gamma_1 \cup \Gamma_2 = \partial\Omega$, $\Gamma_1 \cap \Gamma_2 = \emptyset$. We further assume that there exist a constant $K > 0$ such that if $\phi_- = 0$, $\lambda_+ = 0$ for a time interval larger than K then $\phi = \lambda = 0$.

The necessary conditions for the above optimization problem are

$$\begin{aligned} \sum_{j=1}^d A_j \frac{\partial \phi}{\partial x_j} &= 0 & \Omega \\ - \sum_{j=1}^d A_j \frac{\partial \lambda}{\partial x_j} &= 0 & \Omega \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= u & \Gamma_1 \\ (\mathbf{A} \cdot \mathbf{n})_- \lambda_- + \eta \phi &= \eta g & \Gamma_1 \\ \lambda_+ + \sigma \eta u &= 0 & \Gamma_1 \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= 0 & \Gamma_2 \\ (\mathbf{A} \cdot \mathbf{n})_- \lambda_- &= 0 & \Gamma_2 \end{aligned} \quad (28)$$

where η is an arbitrary positive number. We use it to derive the convergence estimate. Consider the following time embedding

$$\begin{aligned} \frac{d}{dt} \phi + \sum_{j=1}^d A_j \frac{\partial \phi}{\partial x_j} &= 0 & \Omega \\ \frac{d}{dt} \lambda - \sum_{j=1}^d A_j \frac{\partial \lambda}{\partial x_j} &= 0 & \Omega \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= u & \Gamma_1 \\ (\mathbf{A} \cdot \mathbf{n})_- \lambda_- + \eta \phi &= \eta g & \Gamma_1 \\ \lambda_+ + \sigma \eta u &= 0 & \Gamma_1 \\ (\mathbf{A} \cdot \mathbf{n})_+ \phi_+ &= 0 & \Gamma_2 \\ (\mathbf{A} \cdot \mathbf{n})_- \lambda_- &= 0 & \Gamma_2 \end{aligned} \quad (29)$$

For analysis of the behavior of the errors we take $g = 0$ and using integration by parts we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} (\|\phi\|^2 + \|\lambda\|^2) &= \langle (\mathbf{A} \cdot \mathbf{n})_+ \phi_+, \phi_+ \rangle + \langle (\mathbf{A} \cdot \mathbf{n})_- \phi_-, \phi_- \rangle - \\ &\quad \langle (\mathbf{A} \cdot \mathbf{n})_+ \lambda_+, \lambda_+ \rangle - \langle (\mathbf{A} \cdot \mathbf{n})_- \lambda_-, \lambda_- \rangle \end{aligned} \quad (30)$$

where the norms denote L_2 norms on Ω and $\langle \cdot, \cdot \rangle$ denotes inner products on the boundary $\partial\Omega$. Eliminating the u from the boundary condition we obtain the following boundary condition that must be satisfied by ϕ, λ

$$(\mathbf{A} \cdot \mathbf{n})_- \lambda_- + \eta \phi_- = 0 \quad \Gamma_1 \quad (31)$$

$$\lambda_+ + \eta \sigma (\mathbf{A} \cdot \mathbf{n})^{-1} \phi_+ = 0 \quad \Gamma_1 \quad (32)$$

Substituting these into the energy estimate we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} (\|\phi\|^2 + \|\lambda\|^2) = & \langle (\mathbf{A} \cdot \mathbf{n})_- - (\mathbf{A} \cdot \mathbf{n})_-^{-1} \phi_-, \phi_- \rangle_{\Gamma_1} - \\ & \langle (\mathbf{A} \cdot \mathbf{n})_+ - \frac{1}{\sigma} (\mathbf{A} \cdot \mathbf{n})_+^{-1} \lambda_+, \lambda_+ \rangle_{\Gamma_1} \\ & \langle (\mathbf{A} \cdot \mathbf{n})_- \phi_-, \phi_- \rangle_{\Gamma_2} - \langle (\mathbf{A} \cdot \mathbf{n})_+ \lambda_+, \lambda_+ \rangle_{\Gamma_2} \end{aligned} \quad (33)$$

The conditions

$$\begin{aligned} (\mathbf{A} \cdot \mathbf{n})_- - \eta^2 (\mathbf{A} \cdot \mathbf{n})_-^{-1} &\leq 0 \\ (\mathbf{A} \cdot \mathbf{n})_+ - \frac{1}{(\eta\sigma)^2} (\mathbf{A} \cdot \mathbf{n})_+^{-1} &\geq 0 \end{aligned} \quad (34)$$

imply that the changes in energy are non-positive. That is, it is either decreasing or stabilized. Stabilization of the energy can occur only for the value zero, since otherwise it means that there exists a non zero solution to the evolution equation such that ϕ_- and λ_+ are zero for all times. This is in contradiction to the assumption about the constraint PDE.

Since η was arbitrary in this analysis, we can choose it small enough so that the first condition holds. Then if σ is large enough the second condition will hold as well. Thus, we obtain convergence if σ is not too small.

5 Numerical Results

In this section we demonstrate the effectiveness of the methods discussed here for an optimization problem governed by inviscid incompressible flow. Let $\Omega = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\}$, $\Gamma_1 = \{(x, 0) | 0 \leq x \leq 1\}$, $\Gamma_2 = \{(x, 1) | 0 \leq x \leq 1\}$

The minimization problem is given by

$$\min_u \frac{1}{2} \int_{\Gamma_1} (\phi - \phi_0)^2 dx + \frac{1}{2} \eta \int_{\Gamma_1} u^2 dx \quad (35)$$

subject to

$$\begin{aligned} \Delta \phi &= 0 & \Omega \\ \frac{\partial \phi}{\partial n} &= u & \Gamma_1 \\ \phi &= g(x) & \Gamma_2 \end{aligned} \quad (36)$$

and all functions are assumed to be periodic in the x direction.

5.1 Finite Dimensional Design Space

We assume that u has the form

$$\sum_{j=1}^q \alpha_j f_j(x) \quad (37)$$

where α_j are constants to be determined and $f_j(x)$ are prescribed functions. The necessary conditions for this problem are easily derived and we use the following time embedding as a solution process

$$\begin{aligned} \frac{d}{dt}\phi - \Delta\phi &= 0 & \Omega \\ \frac{d}{dt}\lambda - \Delta\lambda &= & \Omega \\ \frac{\partial\phi}{\partial n} &= u & \Gamma_1 \\ -\frac{\partial\lambda}{\partial n} + \phi &= \phi_0 & \Gamma_1 \\ \lambda &= 0 & \Gamma_1 \\ \phi &= g & \Gamma_2 \\ \int_0^1 f_j(x)\lambda(x,0)dx + \eta\alpha_j &= 0 & j = 1, \dots, q \end{aligned} \quad (38)$$

This time dependent process was approximated by Jacobi relaxation, where at each time step all boundary conditions are satisfied. Residuals history is given in fig 2 and shows that the convergence rate is independent of the number of design variables.

5.2 Infinite Dimensional Design Space

In this case we look for u in a function space, namely, $L_2(0,1)$. The necessary conditions are stationary solution of the following time evolution equation which was used in the computation.

$$\begin{aligned} \frac{d}{dt}\phi - \Delta\phi &= 0 & \Omega \\ \frac{d}{dt}\lambda - \Delta\lambda &= 0 & \Omega \\ \frac{\partial\phi}{\partial n} &= u & \Gamma_1 \\ \phi &= g(x) & \Gamma_1 \\ -\frac{\partial\lambda}{\partial n} + \phi &= \phi_0 & \Gamma_1 \\ \lambda(x,0) + \eta u(x) &= 0 & \Gamma_1 \\ \lambda &= 0 & \Gamma_2 \end{aligned} \quad (39)$$

This time dependent process was approximated by Jacobi relaxation, where at each time step all boundary conditions are satisfied. Residuals history is given in fig 3 and shows that the convergence rate is independent of η . It can be seen from that figure that the number of iteration for the analysis problem and for the full optimization problem are not much different.

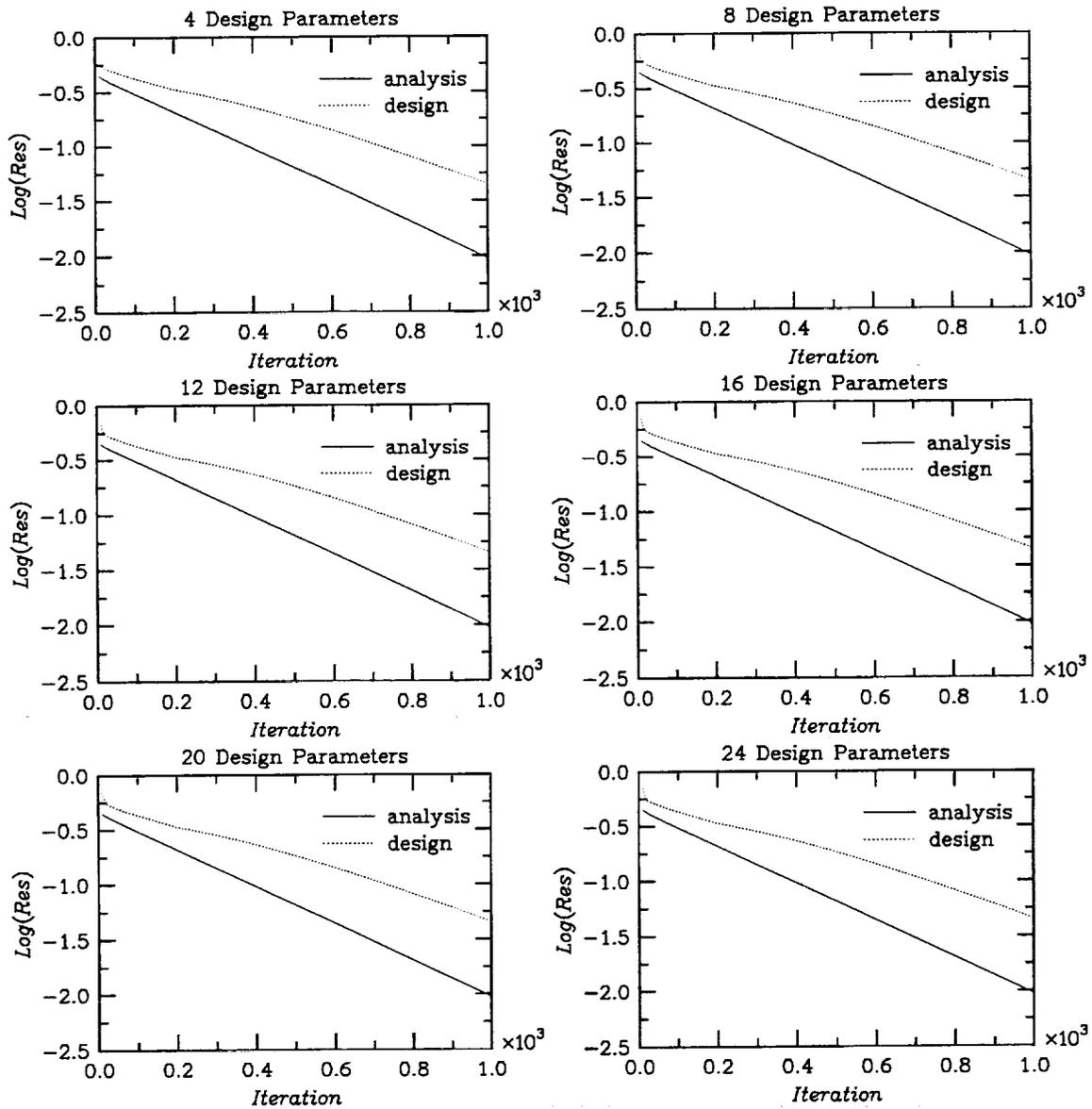


Figure 2: Residuals History for analysis and full optimization method, grid 32×32 , $\eta = 0$.

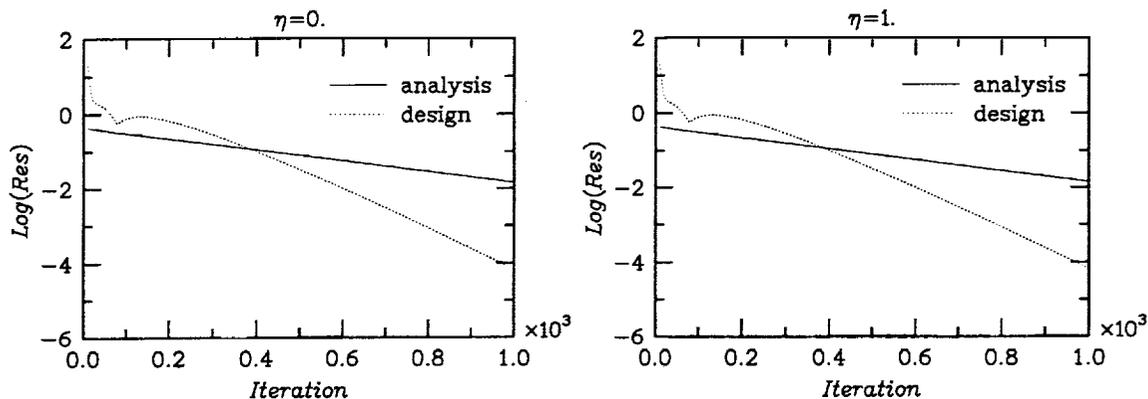


Figure 3: Residuals History for analysis and full optimization method, grid 32×32

6 Conclusions

In this paper we have introduced pseudo-time methods for the efficient solution of optimization problems governed by partial differential equations. In these methods the marching toward the solution of the optimization problem is done on the design hypersurface rather than the intersection of the hypersurfaces for state and costate equations. Very efficient solvers are obtained as indicated from the proofs as well as from the numerical examples included. The methods allow the solution of the full optimization problem in a computational cost similar to that of solving the constrained PDE. The methods do not require gradient calculation however, it is essential to use it with the adjoint equations. The methods offer an alternative to gradient descent methods. Their implementation is straightforward and can be done using multigrid algorithms or single grid iteration.

Acknowledgment The author wishes to thank David Kinderlehrer for valuable discussions.

References

- [AT1] Arian, S. Ta'asan: Shape Optimization in One Shot. Proceedings of a Workshop on Optimal Design and Control held at Blacksburg, VA, April 8-9, 1994.
- [AT2] Arian, S. Ta'asan: Multigrid One Shot Methods for Optimal Control Problems: Infinite Dimensional Control. ICASE Report No. 94-52.
- [BD] F. Beaux and A. Dervieux: A Hierarchical Approach for Shape Optimization, Inria Rapports de Recherche, N. 1868 (1993)
- [CGBN] A. Carle, L.L. Green, C.H. Bischof and P. Newman: "Application of Automatic Differentiation in CFD" AIAA 94-2197, June 1994.

- [GNH] L. Green, P. Newman and K. Haigler: "Sensitivity Derivatives for advanced CFD Algorithms and Viscous Modeling Parameters via Automatic Differentiation", AIAA 93-3321, 1993.
- [HMY] W.P. Huffman, R.G. Melvin, D.P. Young, F.T. Johnson , J.E. Bussoletti, M.B. Bieterman and C.L. Hilmes: "Practical Design and Optimization in Computational Fluid Dynamics", AIAA 93-3111 July 1993.
- [J] A. Jameson: "Aerodynamics Design via Control Theory" J. Sci. Comp. Nov. 21, 1988.
- [T1] S. Ta'asan: One-Shot Method for Optimal Control of Distributed Parameter Systems I: Finite Dimensional Control. ICASE Report No. 91-2, 1991.
- [TSK1] S. Ta'asan, M.D. Salas and G. Kuruvila: Aerodynamics Design and Optimization in "One Shot". Proceedings of the AIAA 30th Aerospace Sciences Meeting and Exhibit, Jan 6-9, 1992, Reno NV.
- [TSK2] S. Ta'asan, M.D. Salas and G. Kuruvila: A New Approach to Aerodynamics Optimization. Proceeding of the First European Conference on Numerical Methods in Engineering, Sept. 1992, Brussels, Belgium.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 1995	3. REPORT TYPE AND DATES COVERED Contractor Report		
4. TITLE AND SUBTITLE PSEUDO-TIME METHODS FOR CONSTRAINED OPTIMIZATION PROBLEMS GOVERNED BY PDE			5. FUNDING NUMBERS C NAS1-19480 WU 505-90-52-01	
6. AUTHOR(S) Shlomo Ta'asan				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23681-0001			8. PERFORMING ORGANIZATION REPORT NUMBER ICASE Report No. 95-32	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Langley Research Center Hampton, VA 23681-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA CR-195081 ICASE Report No. 95-32	
11. SUPPLEMENTARY NOTES Langley Technical Monitor: Dennis M. Bushnell Final Report To be submitted to SIAM Journal on Optimization				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified-Unlimited Subject Category 64			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) In this paper we present a novel method for solving optimization problems governed by partial differential equations. Existing methods use gradient information in marching toward the minimum, where the constrained PDE is solved once (sometimes only approximately) per each optimization step. Such methods can be viewed as a marching techniques on the intersection of the state and costate hypersurfaces while improving the residuals of the design equation per each iteration. In contrast, the method presented here march on the design hypersurface and at each iteration improve the residuals of the state and costate equations. The new method is usually much less expensive per iteration step, since in most problems of practical interest the design equation involves much less unknowns that that of either the state or costate equations. Convergence is shown using energy estimates for the evolution equations governing the iterative process. Numerical tests shows that the new method allows the solution of the optimization problem in a cost of solving the analysis problem just a few times, independent of the number of design parameters. The method can be applied using single grid iterations as well as with multigrid solvers.				
14. SUBJECT TERMS Constrained optimization; pseudo-time methods; non-gradient methods			15. NUMBER OF PAGES 15	
			16. PRICE CODE A03	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	